Performance, Portability and Productivity

"The Three P's"

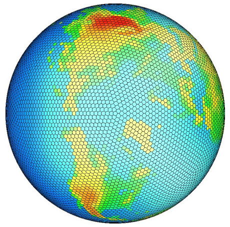
ESiWACE2 has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 823988

esiwace
CENTRE OF EXCELLENCE IN SIMULATION OF WEATHER AND CLIMATE IN EUROPE

Overview

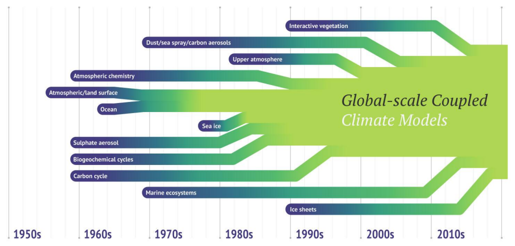
ESiWACE2 has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 823988

Performance



Climate models

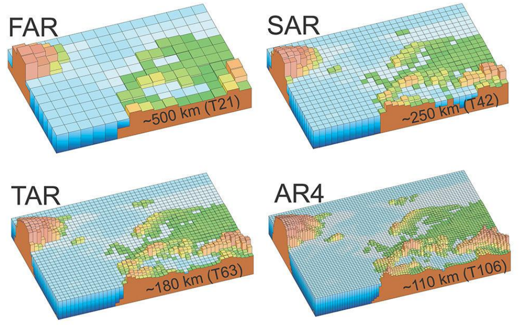
For decades scientists have been using mathematical models to help us learn more about the Earth's climate. Known as climate models, they are driven by the fundamental physics of the atmosphere and oceans, and the cycling of chemicals between living things and their environment. Over time they have increased in complexity, as separate components have merged to form coupled systems.



Note: There were some very simplified models before the dates mentioned.

CarbonBrief

- Why?
- Grand challenge problem - always need more resources
  - Increase resolution, increase time period, increase complexity of science
- Weather/climate has massive political and economic implications so funds available
- Use the biggest machines - very expensive - must get the best out of them - science per second
- Currently aiming for model with 1km resolution
  - ~0.5 billion squares, ~50 billion squashed cubes



ESiWACE2 has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 823988

esiwace CENTRE OF EXCELLENCE IN SIMULATION OF WEATHER AND CLIMATE IN EUROPE

Performance - example

- Met Office currently uses three Cray XC40's
  - ~0.5M CPU cores in total
  - Electricity bill of hundreds of thousands of £/yr
- Used for climate prediction, weather prediction, ocean modelling etc. etc.
- £1.2bn allocated for MO HPC over next ten years

ESiWACE2 has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 8

(Performance) Portability

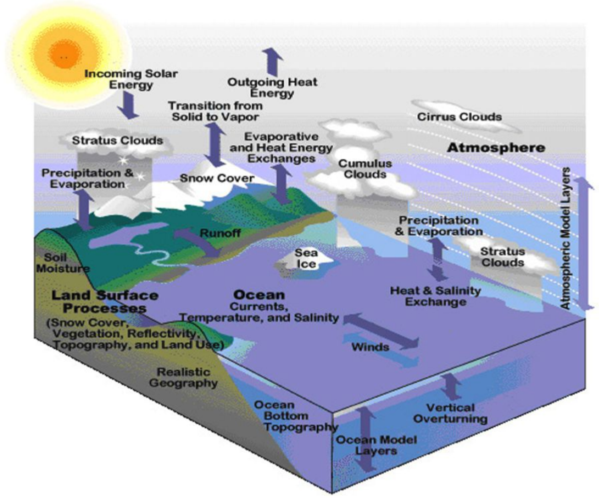
- Why?
- If just one institution and machine then could simply optimise code for that

However...

- Models are used by various institutions who will use different machines, different compilers and different model configurations
- Models are developed over decades, supercomputers tend to be upgraded far more frequently
- If limited to using a particular machine or class of machine then might not get best deal when buying
  - More machine → more science.

ESiWACE2 has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 823988

Productivity



- Very large, (scientifically) complex code
  - UM has ~2M lines of code
- Code continually being developed (over decades)
- Many contributors, often from multiple institutions
  - e.g. LFRic has 10-20 scientists developing code as well as the core LFRic team of ~5-10 scientific software engineers
- Need to keep development time down - avoid bugs, readable code
- Code needs to be correct
  - Used for safety-critical forecasts and political decision making

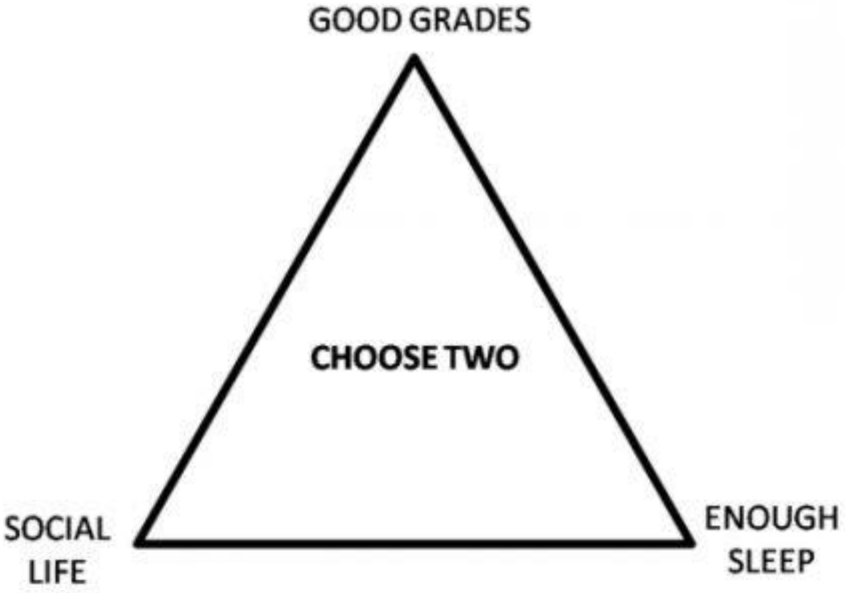
ESiWACE2 has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 823988

Productivity

- Domain scientists may not be performance experts
- HPC/Optimisation experts may not be domain scientists
- Limited pool of people (can't compete with industrial salaries)
- Software engineering problem - s/w development practices - repos, version control, code standards, code reviewing, test suites

ESiWACE2 has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 823988

What is the problem?

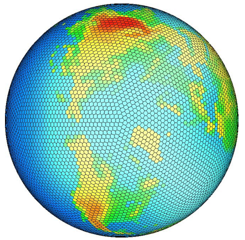


Conflicting goals

- To get fast code you may need to break good software engineering practices
  - Optimised code can be hard to understand and hard to extend

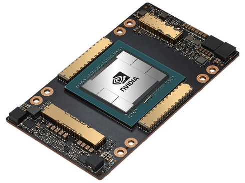
Example:

- Good Software Engineering Practice: "Isolate individual steps of an algorithm into functions"
- Optimizing Technique: "Avoid function calls since setting up a stack frame imposes runtime penalty"
- Main function of dynamical core in ICON is ~3000 lines
  - Example: Tridiagonal Solver pasted right into function body



ESiWACE2 has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 823988

Conflicting goals

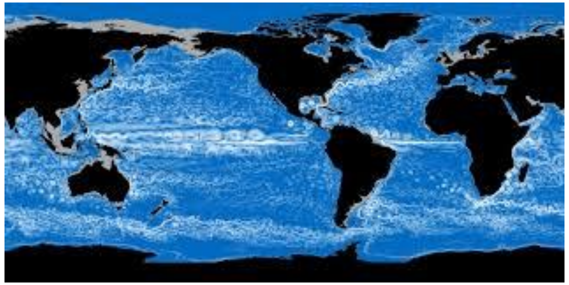


- Optimised code may be different for different architectures
  - Performance portability may require multiple versions of source code

- Model code typically riddled with conditional compilation and other directives
- Hard to test, coverage different depending on compiler flags
- Taking the main function of ICON dycore as an example again:
  - 46 #ifdefs for conditional compilation
  - 426 annotations to support both OpenACC and OpenMP

ESiWACE2 has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 823988

What do institutions do?



- Make trade offs between the 3Ps
    - Details depend on the institution
    - e.g. the Met Office traditionally emphasises scientific performance
    - e.g. NEMO Consortium emphasises scientific performance and portability
- NEMO only runs with MPI or in serial
    - Contains no directives - will not run on GPUs
- ICON runs on CPU and GPU but multiple directives required affecting maintenance. Performance is good, not necessarily highest
- Not many models that run on GPUs
- Optimise for own architectures
- Make optimisations configurable where possible e.g. blocking

ESiWACE2 has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 823988

Summary

- Weather and Climate models are large and complex, are developed by large teams of people, are used by multiple institutions and need to be able to run on some of the biggest supercomputers
- We therefore need Performance, Portability and Productivity
- However, these aims conflict with each other
- Currently institutions make trade-offs between the 3P's
- Productivity (Science development) is usually the strongest driver

ESiWACE2 has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 823988

Next

- Introduction to DSLs
- Break
- Dusk and Dawn
- PSyclone
- Tutorial

ESiWACE2 has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 823988